DOCUMENT RESUME

ED 385 709 CE 069 594

AUTHOR Campbell, Michael D.
TITLE A Review of the Oregon and Texas Experience in Building Performance Measurement and Reporting Systems: Data Selection, Collection, and Reporting. A Publication of the Performance Measurement Reporting and Improvement System Building Project.
INSTITUTION National Inst. for Literacy, Washington, DC.
PUB DATE Jan 95
NOTE 46p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Accountability; Adult Basic Education; Adult Literacy; Agency Cooperation; Case Studies; Comparative Analysis; Data Collection; Data Processing; Evaluation Criteria; *Information Systems; *Literacy Education; *Outcomes of Education; Performance; Program Evaluation; Shared Resources and Services; State Standards; *Statewide Planning; *Systems Approach
IDENTIFIERS *Oregon; *Texas

ABSTRACT

Texas' and Oregon's experiences in building a performance measurement and reporting information system (PMRIS) for their adult literacy programs were compared. The Oregon PMRIS development effort was a multiyear project involving 10 programs and calling for collection/reporting of data regarding 10 of 272 benchmarks recommended by the Oregon Workforce Quality Council. Efforts to develop the Texas PMRIS focused on measures characterizing the labor market success of Job Opportunities and Basic Skills Program participants. The following steps were among those identified as prerequisites for developing a PMRIS: ensure strong direction and commitment from the top; use externally imposed deadlines to further the process; involve potential audiences for information early in the process; focus on a process of continuous improvement; link performance measures directly to vision, benchmarks, and outcomes; balance the use of existing and new, more precise performance measures; and consider the full implications of standardizing data collection and reporting across agencies and programs. (Appendixes constituting approximately 70% of this document contain information on the following: history/experience of PMRIS development in Oregon and Texas, sample performance measures, and shared information systems in Oregon and Texas.) (MN)

A Publication of the

Performance Measurement Reporting and Improvement System Building Project

A REVIEW OF THE OREGON AND TEXAS EXPERIENCE IN BUILDING PERFORMANCE MEASUREMENT AND REPORTING SYSTEMS:

Data Selection, Collection and Reporting

NATIONAL INSTITUTE FOR LITERACY • 800 CONNECTICUT AVENUE, NW • WASHINGTON, DC 20006

A REVIEW OF THE OREGON AND TEXAS EXPERIENCE IN BUILDING PERFORMANCE MEASUREMENT AND REPORTING SYSTEMS

Data Selection, Collection and Reporting

National Institute for Literacy
January, 1995

PREFACE

This report was prepared for the National Institute for Literacy by Michael D. Campbell under contract with Brizius & Foster. It is intended for use by states that participated in the Academy on Performance Accountability held in Washington, D.C., in December 1993. It may also be useful to other states interested in developing performance measurement, reporting, and improvement systems for literacy.

The author wishes to thank the many individuals in Oregon and Texas who spoke freely and in detail about their states' experiences using collaborative interagency approaches in developing performance measurement systems. In Oregon, these individuals include Dewey Harris, Tim Houchen, Mike Marsh, Holly Miles, John Westine, and Debbie White. Those in Texas include Carol Barron, Jacqlyn Dial, Cherrie Dorman, Jack Eshelman, Lea Isgur, Gerald Kaderli, Sandra Simon, and Deborah Stedman. Their candid responses to all questions and the materials they shared were invaluable to the preparation of this report.

Thanks also to Sondra Stein and Susan Foster for their help in framing the questions that were asked in this study and for their thoughtful critiques of this report.

Finally, this report is dedicated to Jack Brizius who passed away shortly before its publication. The author salutes Jack for his lifetime of work to improve public policy and for so many other invaluable contributions to his colleagues, both professional and personal.

CONTENTS

A REVIEW OF THE OREGON AND TEXAS EXPERIENCE IN BUILDING PERFORMANCE MEASUREMENT AND REPORTING SYSTEMS

Data Selection, Collection and Reporting

INTRODUCTION

In 1993, the National Institute for Literacy awarded grants to five states to undertake the development of performance measurement systems for literacy. In performance-based systems, literacy policy will be judged by evidence of accomplishment of publicly stated goals for improving statewide literacy. Literacy programs will be assessed by the results they achieve for participants. Information from literacy performance measurement systems will be used to improve program management, promote continuous improvement in program operations, and revise and adjust policies. Ideally, the systems for tracking policy and program outcomes will be integrated so that the relative and cumulative effects of various literacy programs on policy outcomes can be gauged.

This is the vision set forth by NIFL when it awarded the grants. The role of the grantee states is to take the first steps toward achieving this vision — to test the waters and lead the way for other states interested in making their literacy systems more results-oriented.

Because no single agency has sole responsibility for improving adult literacy levels in a state, each grantee state assembled a team of high-level officials from the major agencies that provide literacy services. The task for each state team under the grant is to plan and begin development of a performance measurement, reporting and improvement system (PMRIS) for literacy. The system should not only monitor program outcomes for adult learners and track progress on statewide literacy objectives, but it should also report this information in such a way that it can be used to make policy and budgeting decisions and to continuously improve the management and operations of individual literacy programs.

Early in the process, NIFL and its consultants suggested ten steps that a state might take in developing and using a PMRIS for literacy. These steps are as follows:

Ten Suggested Steps in Developing and Using A PMRIS

Step 1: Define Vision and Benchmarks

Step 2: Define Policy Outcomes

Step 3: Identify Performance Measures for Policy Outcomes

Step 4: Define Program Outcomes Related to Policy Outcomes

Step 5: Identify Program Performance Measures

Step 6: Identify Target Populations Related to Policy and

Step 7: Compare Existing Service Delivery Patterns With Needs

a. Determine what program interventions are necessary to achieve outcomes for target populations
b. Determine whether existing programs can achieve these outcomes.
c. Apply quality standards to existing and new programs.

Step 8: Collect Data and Communicate Results

a. Define carefully the audiences
b. Develop sample reports
c. Negotiate data collections protocols
d. Collect data
e. Report results

Step 9: Use Information From PMRIS to Improve Program Interventions

a. Link inputs to processes
b. Identify constraints and enablers
c. Plan for continuous improvement

Step 10: Use Information From PMRIS to Adjust Resources and Policies.

The state teams have worked through the first six steps to various degrees. In this project year, they will be working on the remaining steps and revisiting the work they did last year. Of particular concern for this project year are those issues related to Step 8: Collect data and communicate results. To help inform states on the issues they will face in building data collection and information dissemination systems, the NIFL has commissioned a review of the progress made in this area in two other states: Texas and Oregon.

These states were chosen because they have been working on performance management somewhat longer than most and because they have had experience in interagency efforts to develop performance measurement systems in the area of workforce development. This area is closely related to, and encompasses some aspects of, literacy programming.

The Oregon case study describes a ten-program effort in developing a shared data collection and reporting system that has been ongoing for several years. The Texas example, in contrast, presents a process that was carried out quickly due to legislative time pressures. The

lessons learned in both states and the implications for developing a PMRIS for literacy are summarized below.

LESSONS LEARNED

Workgroup members from both states were asked what advice they might give to other interagency groups contemplating the joint development of performance measurement systems, particularly data collection and communications systems. Many useful lessons emerged. These lessons fall roughly into two categories. The first set concerns the process of developing data collection and communication components of performance measurement and reporting systems. These address the *prerequisites* for getting the job done — things like leadership, timing, and commitment. The second set of lessons concerns the *issues* that interagency teams must address and the decisions they must make in selecting performance measures and establishing reporting systems. These include things like weighing the cost of data collection against the need for valid and reliable data, and the issue of using standard measures across programs. Each is discussed below. Additional background information is found in the appendices.

PREREQUISITES FOR DEVELOPING A PMRIS

1. There must be strong direction and commitment from the top.

All of those interviewed believed that their interagency group would never have accomplished as much as it did — in fact, probably never would have formed — without the support and direction provided by a higher state authority. This lesson has certainly been confirmed in the experience of the states participating in the NIFL process.

In Oregon, this authority was the Workforce Quality Council established by Act of the Legislature in 1991 to ensure, as its strategic plan stated, that "Oregon will have the best educated and prepared workforce in the nation by the year 2000 and a workforce equal to any in the world by 2010." The Council made performance measurement one of its top priorities. The majority of the Council's twenty-one members are from the private sector, but the Council also includes the top state officials who oversee workforce agencies. In all, the Council is responsible for coordinating the efforts of ten state agencies or sub-agency units that are involved in workforce education and training. One of its key charges is to, "Provide cost-effective coordination and accountability of existing employment and training programs."

The Texas team benefited from the legislative mandate and authority for action provided by Rider 33, a rider attached to the 1993 Texas general appropriations act concerning the JOBS program. Rider 33 required the state to develop intermediate outcome measures for all components of the JOBS program. Team members in both states said their ability to get things done was greatly enhanced when there was a strong commitment to the project by agency leaders. As a Texas team member pointed out, the combination of legislative interest and top-level support within agencies allowed them to get the resources they needed and to remove roadblocks quickly.

A state team that wants to institute a PMRIS for literacy would do well to find a "champion" — the governor, a legislator, an influential statewide council, or even an auditor — to push for the development of the system. The team can help the process along by supplying the champion with a solid rationale for why the system is needed. The Oregon team, for example, provided the Council with a list of tangible and intangible benefits that would accrue from a shared information system. The Council subsequently convinced the legislature to appropriate funds for the system, even though the state budget was particularly tight. (For background on the development of a performance management system in Oregon, see Appendix A.)

Related requirements for progress are that the group have an organized, task-oriented leader, members who are committed to the cause, and top-quality support staff in each of the agencies. Workgroup members in both Oregon and Texas cited all of these factors as essential to the success of the work they undertook. The team must not only have these necessary ingredients from the beginning, but it must ensure that they are not lost through team membership changes or turnover in agency leadership during the sometimes lengthy process of developing a PMRIS.

2. Staff and resource commitments are large, and change in nature over time.

Team members in Oregon and Texas found that the interagency process was both time and labor intensive. Identifying appropriate measures and agreeing on standards for each is a challenging, time-consuming task. Not only did they have to meet long hours with the other group members, but they also had to spend a lot of time in their own agencies briefing others, gathering needed information, and moving the interagency agenda. The experience of the NIFL states confirms that interagency team work is particularly intensive during the first year of developing a performance measurement system. Teams in both Oregon and Texas met weekly or every other week for months during the initial phase of their work.

The subsequent tasks of testing, implementing, and refining a performance measurement system require less work from the interagency team members, but more work from technical and support staff within the agencies. This is an issue that NIFL states should keep in mind as they plan their activities and staff support for this and subsequent years. Technical and support staff from within the agencies must receive adequate orientation and in-service training on the nature and purpose of the PMRIS if they are to be productive in helping to institutionalize it. This latter phase can take years, depending on the scope of the work.

While the participating NIFL states have confirmed that the process of developing a performance management system requires a long term commitment, the experience in Oregon and Texas underscores the fact that the second and third years may require a different alignment of resources than the first. The first year is labor intensive on the policy side — designing and selling the system, developing the vision, goals, benchmarks and outcomes, orchestrating the inside and public information processes. Subsequent years will require continued work to refine the initial policy framework, but much more intensive involvement on the part of support staff within the agencies and provider staff in order to implement the system.

3. Use externally imposed deadlines to move the process along.

Dr. Johnson observed that, "Nothing so focuses the mind as knowing that you will be hanged in the morning." In both Texas and Oregon, firm deadlines for gathering performance data and reporting it to the legislature were imposed on the workgroups. These externally imposed deadlines helped keep the teams focused and moving along on their tasks. The downside of having such deadlines is that sometimes teams may be tempted to sacrifice quality for expediency — e.g., choosing to use an inferior, but available, performance measure in order to have "something to present to the legislature in January."

On the whole, deadlines imposed by the legislature, governor, or other external authority, in the experience of Texas and Oregon, proved more helpful than harmful. They helped legitimize the work of the teams and often imposed needed discipline on the process. As one team member remarked, "a performance measurement system will become institutionalized more quickly if a periodic demand for data from the system is built into the regular reporting cycles for the legislature and the governor."

NIFL states may want to consider creating some externally imposed deadlines in cooperation with the governor's office, legislators, the business community or other stakeholders. Deadlines such as these could help provide some state structure and timetable for the development process, rather than just relying on the timetable imposed by the grant. Externally imposed state deadlines will also help to continue the system change effort after the grant has expired.

4. Involve the potential audiences for information early in the process.

The Texas workgroup found that it paid to have representatives from the legislative budget and auditor's offices at the table when they were contemplating what to measure and report. It reduced the potential for misunderstanding what these key audiences wanted, and thus saved everyone time. Early in the process in Texas, the workgroup asked the legislative budget office and the state auditor's office to assign representatives to the workgroup. Agencies got direct guidance from these representatives throughout the process on what information their offices would find most useful.

A major result of this collaboration, one member said, was that the agency workgroup members "didn't waste time trying to figure out what these guys really wanted in response to Rider 33." Another result was that intermediate outcomes (see Appendix B for background on building intermediate outcome measures in Texas) identified by the group eventually were divided into two categories: (1) *outcomes for external reporting purposes*, which were measures needed by the legislative budget and audit offices, and (2) *outcomes for management purposes*, which were measures to be used internally by agencies to manage their components of the JOBS program.

The Texas team also found it useful to have state and regional staff from their agencies comment on the design of the management reports being prepared for them. As one interviewee put it, "If the staff responsible for gathering data are involved in selecting the measures, their sense of ownership increases and they will do a better job of recording information. They will have even

more incentive to record data accurately if they also know they will receive management reports based on their data."

Oregon conducted a study prior to development of its shared information system to determine what information potential users might need. Unlike systems being built in the NIFL states, the Oregon system is not intended to provide program management information. Consequently, regional and local staff were not involved in decisions about what data to gather and report. As a result, there is more concern in Oregon than in Texas about the accuracy of the data collected by the various agencies.

The lesson for the NIFL states is that involvement of stakeholders early in the process, while at times difficult and cumbersome, results in greater investment in and commitment to the system once it becomes operational, and minimizes the chances of significant change down the line. Transitions are difficult. People don't want to give up their old ways of doing things until they are confident the new way will work better. For instance, the first time state agencies in Texas submitted performance-based budgets to the legislature, the response was: "This outcome information is very interesting and useful, but please also submit traditional budgets that show where all the cars and furniture and things are." The same thing happened when the Texas workgroup submitted their list of proposed intermediate outcome measures to the legislative and auditor's staff for review. They responded by asking for some additional input and process measures of the type they were used to getting. Involving stakeholders early in the process can help to minimize these problems.

When involving stakeholders, it is important to try and create a win-win situation. One of the reasons the Texas group worked so well together was that each agency had something to gain from the collaboration. As the prime agency for JOBS, the department of human services and its clients would benefit greatly if the data collection and reporting system helped improve services in other agencies. The other agencies stood to gain better management information from the monthly reports prepared by the human services department. In addition, all agencies could garner more credit from the legislature for their accomplishments with JOBS clients — achievements that previously were not recognized. The key to a win-win situation is a clear understanding on the part of those involved of what the benefits are for them.

The result of involving stakeholders in the process of selecting data items may well be the inclusion of some input and process data that is not really needed for outcome-based decision-making. Trade-offs such as these may have to be made initially, at least until those using the system have become accustomed to it and appreciate its value.

5. Don't expect perfect outcome measures the first time — but focus on a process of continuous improvement.

Workgroup members from both Oregon and Texas readily admit that the outcomes they have identified and the performance measures they are currently using are far from perfect. Oregon, in fact, openly acknowledges this in referring to their performance measures as "interim" measures. Staff in both states believe that the measures they developed are at least serviceable and

sufficient to get the data collection and reporting process geared up and moving. They both plan to improve the measures with each successive cycle of data collection and reporting.

One person interviewed urged PMRIS states to move past their fear of selecting the wrong measures: "If a state developing a PMRIS for literacy wants to wait until it has selected the perfect measures before implementing the data collection and reporting process, it may never get beyond the planning stage."

Experience in both Oregon and Texas, as well as the experience of the NIFL states themselves, has shown that debate about what are the best measures and whether it is feasible to gather the right information can be interminable. By trying out the system, even with less than perfect measures, a state can identify weaknesses in the data collection and reporting process. It can correct these weaknesses at the same time it validates and revises the performance measures.

Both Oregon and Texas stress, however, that it is important to make all parties aware that there must be period of trial and error — of developing and phasing in appropriate outcome and performance measures. Legislators and governors, in particular, should be convinced to delay using information from a PMRIS to make funding decisions until at least a couple rounds of improvements have been made.

ESTABLISHING DATA COLLECTION AND REPORTING SYSTEMS

To collect data and report results, the Oregon and Texas workgroups had to make many critical decisions about what to measure and how to measure it. An interagency team developing a PMRIS for literacy will be forced to confront many of the same issues and choices. The following are some of the lessons on selecting what to measure and report that can be drawn from the experiences in these two states.

6. Link performance measures directly to vision, benchmarks and outcomes.

In selecting performance measures and data elements for the PMRIS, it is easy to become lost in the mire of data already available. Clarity is essential. The desired outcomes in human services and education are changes in people's lives — changes that should occur if a policy or program is working as intended. These can include changes in behavior, circumstances, knowledge, attitudes, skills, and abilities. The changes must be linked to the outcomes and overall vision, goals or benchmarks that you have selected. In designing a PMRIS, experience of Oregon and Texas suggest that it may be helpful to:

- specify which groups of people are served by a program,
- identify the main changes in their lives the program should produce in order to meet the outcome, and
- develop valid measures that indicate whether these changes occur.

The Experience in Oregon: Oregon's Workforce Quality Council tried to make the workforce agencies it oversees more accountable for results by identifying overarching benchmarks to guide the efforts of all agencies. The Oregon Workforce Quality Council culled through the 272 Oregon benchmarks and identified ten benchmarks to be the top priority for the Workforce Quality Council. (These are presented in Appendix C.) The selected benchmarks represent areas that the Council and its member agencies can affect through their programs and policies.

These benchmarks do not actually attempt to measure workforce functioning directly. There is no benchmark, for example, that tracks the percentage of workers rated as highly productive by their employers. Instead, the benchmarks track certain qualities in workers and employers that are usually associated with high productivity: good language skills, post-secondary education or training, professional development opportunities for employees, etc. The selected benchmarks, therefore, are really proxy measures for workforce productivity. The logic is that if these qualities improve, so will productivity and Oregon's workforce will compete successfully in the world economy.

Three of Oregon's benchmarks deal directly with the issue of adult literacy and are measures of the functional literacy skills in English of adult Oregonians, ages 16-65. These benchmarks were developed through a statewide survey in 1990 that measured the prose, document, and quantitative literacy levels of a representative sample of adults. The survey was similar to the National Adult Literacy Survey, which was conducted in 1992. As with the other benchmarks, these literacy measures are proxies for real achievement. A more direct — but also more costly measure — would be the percentage of workers who demonstrate adequate prose, document, and quantitative skills on the job.

The Workforce Quality Council identified a few key changes (or outcomes) it wants to see for every group to achieve these benchmarks: high post-program employment rates, good wages, job retention, and improved functional literacy. The performance measures it currently uses — job and wage data from the unemployment insurance files and improved literacy scores on the CASAS — were selected because they felt they provided standard measures of these changes at a relatively low cost.

While these measures answer the fundamental questions of change for most people served by Oregon's workforce development programs, they do not adequately measure change for every participant. There are gaps in measurement. Some people, for example, get jobs in industries that are exempt from unemployment insurance reporting; consequently, their employment gains are not measured. In addition, the CASAS, a paper-and-pencil test which has been used as a proxy measure for improved functioning in the workplace, is of questionable value in showing a direct link to other outcomes and cannot detect changes for people at the higher end of the literacy continuum. Having identified these gaps, the Oregon team is working on developing better, more comprehensive measures. (See Appendix D for a description of Oregon's Mobility Continuum.)

The Experience in Texas: In the Texas case, JOBS program participants are the target population. The desired program outcomes are that they become employed and permanently leave the welfare rolls. Performance measures for these outcomes (employment, wages, welfare

termination and recidivism) were already established before the interagency workgroup was formed. The group's task is to find ways to measure and report whether participants are making adequate progress in the program. To do so, they must answer this fundamental question for each program component, "What has changed in the lives of people who participated in this component?" Ideally, each component will produce an incremental change in a participant's behavior, knowledge, skills, etc. that should move the person closer to achieving the ultimate program outcomes. (For a more detailed description of performance measures developed in Texas, see Appendix E.)

Reconciling Existing and Needed Data Items: In selecting performance measures, NIFL states will also want to examine those data items currently collected to determine their utility in measuring progress toward outcome achievement. This will involve three steps:

- Identify those data items currently collected that will measure progress toward achievement of outcomes, or quality improvement.
- Identify those data items currently collected which are not needed for purposes of outcome measurement or quality improvement, but which are required by federal or state law.
- Identify those data items currently collected which are not needed for either purpose.

Those in the last category can be considered for elimination. Those in the second category can be the subject of debate with other state or federal officials to consider changes in state law and waivers of federal requirements. By eliminating unnecessary data items, the current reporting burden on providers could be reduced.

7. Balance the use of existing performance measures and developing new, more precise measures.

Both interagency teams weighed a number of factors in deciding what measures they would use to monitor program outcomes. They knew that a perfect measure would be precise, valid, reliable, meaningful, and useful for decision making. Ideally, it would also be inexpensive to gather and already available somewhere in the system. Needless to say, they found no measures that met all of these criteria. There were always tradeoffs.

One of the most frequent decisions the teams faced was whether to develop a new measure that would actually measure the outcome of interest or to rely on an existing measure that "looked close enough." Should they, for example, try to measure whether program participants were actually more productive as a result of skills training, or should they use readily available data such as number of hours of training or test score gains, as evidence? The arguments for going with existing measures are many: no development and training costs, people are used to gathering and using the data, and data are available quickly — an important consideration when a deadline for a report to the legislature looms. But, if the existing measures don't produce data that answer the questions of "Are people making progress, changing their behavior, etc.?" then using them just because they are available doesn't make sense.

In many instances, the interagency teams in Oregon and Texas chose to use imperfect, existing measures as proxies for what they really wanted to measure. Usually, this was because they didn't have the time or resources needed to develop more exact measures. Both teams, however, openly acknowledge the problems in using these measures. They are seeking additional resources to develop better measures for future rounds of measurement and reporting.

In making the decision whether to use existing measures, it is important for PMRIS states to honestly evaluate the relationship between the measure selected and the outcome to be achieved. For example, if state officials choose to use seat time as a proxy for productivity improvement, they would be kidding themselves that this tells them anything about progress toward outcome achievement. Putting existing but useless measures into a new format and process will result in the stillbirth of the PMRIS. (For more detailed descriptions of the experiences of Oregon and Texas in developing performance measures, see Appendices F and G.)

8. Consider the full implications of standardizing data collection and reporting across agencies and programs.

The Oregon and Texas case studies raise a number of issues about standardizing data collection and reporting. The Oregon workgroup attempted to develop standard outcome measures to be used by all workforce development agencies. Oregon's Workforce Quality Council, in their review of the budget process for workforce programs found that federal, state, and local agencies use a maze of different computer systems to accumulate data on these programs, and there is no system to integrate information so that the Council could determine where the money is going and what results are achieved for individuals completing education and training programs. In response to these findings, the Council tried to make the workforce agencies it oversees more accountable by identifying common measures that could be used to measure participant outcomes regardless of which agency provided the service, and by establishing a system for sharing information on participants and their outcomes across agencies.

Oregon quickly found that they had to standardize more than just their outcome measures. To avoid unfair comparisons when looking at performance across programs, they also had to standardize many other data elements as well (client characteristics, program costs, etc.). This led to the development of a 120-page data dictionary and the realization that each agency would need to add about 15 to 20 new data elements to its current data collection system. The costs of making such major changes in their data collection systems — in terms of forms revision, computer programming, staff training, etc. — were prohibitive for many of the agencies. In addition, some agencies were reluctant to make these costly changes to their data collection systems because they didn't believe the standardized measures would give them the program management information they need.

Texas sought to avoid some of these problems by allowing each agency to define its own intermediate outcome measures and to simply report "yes" or "no" whether a client was making adequate progress. This minimized the cost of making data systems changes for them and allowed agencies to develop measures that would be useful for program management and promoting continuous improvement in program operations. The problems with this approach are that it is impossible to make comparisons across programs that ostensibly are serving the same clients and

striving for the same outcomes, and, it is impossible to quantify progress toward outcome achievement. A "yes" in one program does not necessarily mean that a client made the same progress as a client receiving a "yes" in another program. In addition, this approach means that each agency must separately validate the measures it uses, rather than sharing the cost and responsibility of validation across agencies.

To link the client records of any individual across agencies requires that the agencies use the same information (name and/or a unique number) to identify case records. The committee in Oregon decided to use participants' social security numbers as the common identifier for all agencies. This presented problems, however, in that some federal programs prohibit using the social security number in this way without safeguards to protect the privacy of program participants.

Oregon's working group sought a legal opinion on the matter from the state's attorney general. The opinion, which took nearly a year to be rendered, prohibited the agencies from using social security numbers in this way until such time as the agencies incorporate an approved information release form in their intake procedures. By signing such a form, participants would allow the agencies to use this number as a means to match records in other agencies — for example, to track post-program wages and job retention using the unemployment insurance tax files. The process of seeking an opinion, developing informed consent forms, and assuring that they are in compliance with standards set forth in the attorney general's opinion created significant delays in the development of the SIS. (For more detailed information on the process of building shared information systems in Oregon and Texas, see Appendices H and I.)

SUMMARY

The experience in the states of Oregon and Texas both confirm many of the issues which have surfaced in the NIFL states working to develop performance management systems and raise additional issues that these states will face in the next year. A summary of these lessons is:

- There must be strong direction and commitment from the top as well as organized, task-oriented leaders, members who are committed to the cause, and top-quality support staff in each of the participating agencies. These resource requirements are essential not only for the initial policy-development process, but also for the long and arduous tasks involved in implementation across agency lines. Staff and resource commitments are large and change in nature over time. The policy work required in the initial phases gives way to the technical and support functions essential to the implementation process.

- The use of externally imposed deadlines — by the governor's office, the legislature, the business community or other stakeholders — can be very helpful to states in moving the process along. These may be of particular value once the NIFL grants expire.

- Involve the potential audiences for information early in the process so as to reduce surprises and changes later on and to build commitment to and investment in the

PMRIS. Those involved in helping to design the system will be allies in its implementation and will be particularly helpful in the implementation phase.

- Don't expect perfect outcome measures the first time — but focus on a process of continuous improvement. Building the PMRIS involves balancing the needs for accurate and useful performance measures and getting the system up and running. It may be necessary to go with less than perfect measures initially, but it is important to keep in mind the need to improve over time. Link performance measures directly to vision, benchmarks and outcomes. These are the standards which drive the system and will help you decide the value of performance measures. Balance the use of existing performance measures and developing new, more precise measures. When using existing measures, however, be sure that they provide information on outcome achievement or performance improvement. Consider eliminating measures that do not provide this type of information.

- Consider the full implications of standardizing state data collection and reporting across programs. In building a PMRIS, states are usually aware of the need to develop performance measures that can be used across programs. In order to report progress, however, states will have to standardize many data collection and reporting elements and processes in order to be able to make comparisons of progress and cost across programs. This will be costly, time consuming, and extremely complicated given the different federal and state reporting requirements now in effect. In order to make progress in this area, states will have to involve providers and other stakeholders in this process and develop a long term strategy to phase in standardization over time.

APPENDIX A

BACKGROUND ON OREGON'S PROGRESS IN BUILDING A BETTER WORKFORCE: FROM BENCHMARKS TO PROGRAM PERFORMANCE MEASURES

"What gets measured gets done, and what gets recognized gets done best." This appears to be the unofficial state motto for Oregon, which has been an acknowledged leader in the field of performance measurement since the publication of *Oregon Benchmarks* in 1991. In that report to the state legislature, the Oregon Progress Board recommended 158 measures (later adopted and increased to 272 measures) that could be used as benchmarks to gauge progress toward a strategic vision of what the state should look like in the year 2010. The vision itself — developed through a series of public forums — had been set forth two years earlier in another report, *Oregon Shines*. This strategic planning document proposed that, through the collective efforts of government, business, and citizen groups, Oregon over the next two decades could achieve sustained economic prosperity, enhance its already high quality of life, maintain its natural environment, and build communities on a human scale.

While other states had created similar, lofty visions of their own futures, none had been able to move effectively from the act of creating a vision of a preferred future to the systematic enactment of public policies and programs that would help create that future. Policymakers, program planners, and the public usually lost sight of their "vision for tomorrow" when they resumed the task of grappling with the immediate problems and issues of the day.

Oregon, in contrast, decided to bridge the gap between vision and action by creating a set of measures — the Oregon benchmarks — that would indicate whether the state was making progress toward the visionary goals set forth in its strategic plan. The priority was on measuring results (for example, the number of adults who are literate) rather than efforts (the amount of money spent on literacy education). The logic was that, by staying focused on outcomes and by keeping track of results, state leaders could reset public priorities and adapt and modify programs as they learned what works. Further, it was decided that this should be a very public process in which progress toward the desired outcomes is measured and reported publicly every two years at the beginning of the biennial legislative session.

Since the adoption of the benchmarks, the ethos of performance measurement has permeated every level of state government in Oregon. The governor and legislature use the benchmarks as a means for setting policy and budget priorities. Some benchmarks identified as critical to the overall success of the state's strategic plan have been given priority attention in terms of resources and position on the policy agenda. In addition, a 1993 act directs all state agencies to develop mission statements and performance measures "consistent with and aimed at achieving Oregon benchmarks." Under this act, top agency managers are responsible for developing the agency mission statement and performance measures with input from all agency employees. Middle management and line staff are responsible for determining the duties they must perform to

achieve the mission. The plan is to use these agency-generated measures in several ways: as a management tool to improve program performance, as a communication tool to inform elected officials and the public about agencies' accomplishments, and as a budgeting tool to allocate scarce resources to programs and units that can demonstrate the best return on investment in terms of results.

One portion of the state strategic plan addresses the need to have a workforce that can compete effectively in a global economy. Accordingly, Oregon intends to improve the overall education and skill levels of its already capable workforce and make it into a world-class workforce by 2010. This intention has resulted in the development of benchmarks for measuring workforce quality. Since a large portion of the workforce in 2010 will consist of adults who today are already out of school and in the labor force, the level of adult literacy is of special concern. Consequently, some of the "core benchmarks" (defined as "fundamental, enduring measures of Oregon's vitality and health") consist of measures of adult literacy.

Oregon's experiences in measuring policy and program performance should be instructive to states that are developing adult literacy performance measurement, reporting, and improvement systems. Of particular interest are:

- How the visioning and benchmarking process in Oregon has influenced the development of performance measures in the area of adult literacy;
- How adult literacy measurement and reporting fits within the overall mission of the state's Workforce Quality Council;
- How the council is attempting to develop a shared information system for state agencies; and
- How all of this work has prompted the creation of a classification model, called the "mobility continuum," which attempts to refine the process of measuring and reporting outcomes for adult learners and others participating in workforce development programs.

APPENDIX B

BACKGROUND ON BUILDING INTERMEDIATE OUTCOME MEASURES IN TEXAS THROUGH INTERAGENCY COOPERATION

In the current movement toward greater governmental accountability, no state has moved faster or done more than Texas to link budget decision-making to the measurement of program results. Texas government crossed the Rubicon in this regard when it passed legislation in 1991 directing state agencies to switch from a traditional zero-based budgeting system to "performance budgeting." Performance budgeting requires each state agency to develop an outcome-oriented strategic plan and to use this plan as the basis for its biennial appropriations request. The plan must contain, among other elements, a statement of the agency's mission and philosophy, its goals and objectives, and the strategies or methods by which it will achieve its objectives. In addition, the agency must develop outcome measures for every objective and report on these outcomes annually. The switch in budgeting methods has, in effect, required Texas state agencies to establish performance measurement and reporting systems for all of their programs.

The process that each state agency uses to produce its performance budget is similar in many respects to the process that states with PMRIS grants from the National Institute for Literacy have been asked to go through. Each agency in essence establishes its own vision and benchmarks, determines what policy and program outcomes it wants to achieve, and establishes a system for measuring and reporting results.

Because the new budgeting system in Texas still tends to focus on the performance of each agency separately, it does not usually cause agencies to work collaboratively in developing outcome measures or data reporting systems — even when the agencies have similar objectives and are serving the same clients. It did, however, indirectly promote one such interagency effort that should be of interest to states that have PMRIS grants. The RIDER 33 workgroup, a team representing various agencies responsible for preparing welfare recipients for the world of work, jointly developed a data collection and reporting system to measure program performance across agencies.

The Rider 33 Workgroup was so named for a rider attached to the 1993 Texas general appropriations act concerning the JOBS program, a state and federally funded job training and education program for welfare recipients in which services are provided by a consortium of state agencies. The human services agency, for example, provides case management services for JOBS clients, the state employment commission and the department of commerce provide job preparation and job training services, and the state education agency and the higher education coordinating board operate the educational components of the JOBS program.

Rider 33 came about in part because the performance budgeting system in Texas created new procedures for monitoring agency performance. Agencies must now submit annual reports to the Legislative Budget Office that report on the outcomes of their efforts (e.g., the percentage of JOBS clients who became employed). They also submit quarterly reports on program outputs (e.g., the number of people served by the JOBS program) and on program efficiency (e.g., the

average cost per client served). The State Auditor's Office—in addition to retaining its traditional role as the fiscal watchdog for state agencies—is now responsible for certifying the accuracy of the data in these reports. It does so by conducting periodic audits of agencies or programs.

An audit of the JOBS program in 1992 found that the outcome measures reported by the various agencies involved in JOBS were not sufficiently detailed to provide the evidence of accomplishment needed by legislative budget staff and other oversight entities. For example, the state education agency might report that 25 percent of the JOBS clients who were enrolled in adult secondary education earned a high school equivalency certificate or a high school diploma during the fiscal year (the standard reporting period). Relying on this statistic alone could lead one to underestimate the accomplishments of the program. This is because students who made substantial educational gains during the year, perhaps improving their educational levels several grade levels, but who did not complete the requirements for a certificate or diploma by year's end, are not counted. Thus the progress (or lack of progress) of 75 percent of the enrollees is not accounted for when just the one outcome measure is reported.

To remedy this type of situation, Rider 33 required the state agencies to develop *intermediate* outcome measures for the various JOBS components. Intermediate outcome measures document the incremental gains made by clients as they progress through a program. In some cases, there are natural transition points in the program at which progress can be monitored. One can measure, for example, the percent of JOBS clients enrolled in beginning-level adult basic education who complete that level and move to the intermediate level. In other cases, client skill levels may be assessed at enrollment and again at a later point in the program to see if progress is being made. The percentage of clients who show improved skills in this test-retest situation would constitute the intermediate outcome measure for the program.

Rider 33 had two other requirements. First, that the intermediate outcome measures from all agencies were to be reported to the Texas Department of Human Services, the coordinating agency for JOBS, and entered into its automated JOBS data system. Second, that these measures be developed quickly. The agencies were given about ten months — from June 1993 until April 1, 1994 — to have all the measures ready for review by the State Auditor's Office and the Legislative Budget Office. Implementation of the measures following the review was to begin September 1, 1994.

To accomplish the tasks set out in Rider 33 in the time period allowed, an interagency workgroup was established composed of program directors, program specialists, and evaluation supervisors from the five state agencies responsible for the various components of the JOBS program. Workgroup participants who were interviewed for this study describe the process of responding to Rider 33 as "time and labor intensive." The workgroup met frequently — at times almost weekly — for over a year. In addition to the workgroup members themselves, the initiative required the services of a variety of personnel in each of the participating agencies. Among these were:

- *program specialists* who reviewed program operations looking for strategic points at which to measure progress, and who helped judge the usefulness of the proposed measures and reports for managing programs;

- *regional and field staff* who provided feedback on the proposed measures, both in terms of their usefulness for improving local program operations and the potential burden the added data collection responsibilities would place on line workers;

- *agency or program directors* who reviewed and approved the proposed measures and reports and authorized the expenditure of agency resources to gather the necessary data;

- *management information system personnel* who reworked data collection forms and computer programs in their respective agencies to handle the new data required by Rider 33, and who created or revised documents for reporting the data;

- *quality assurance specialists* who trained program staff in how to gather and record data, and who developed procedures for sampling case records to review the accuracy and completeness of the data;

- *evaluators* who developed and field tested new instruments to measure progress for several of the components—for example, developing a self-inventory of job search skills that could be administered at two or more points during job preparation training; and
- *budget officers* who reviewed proposed outcomes and measures to ensure that they meshed with the state's performance budgeting process.

The agencies represented and their areas of responsibility were:

- Texas Department of Human Services (lead agency)
  - Case management
  - Job readiness: Life skills

- Texas Education Agency
  - Adult secondary education
  - Adult basic education
  - English as a second language

- Texas Employment Commission
  - Job preparation
  - Job readiness: Life skills for employability enhancement
  - Individual/group job search

- Texas Department of Commerce
  - Occupational skills training and all other JTPA activities
  - On-the-job training

- Texas Higher Education Coordinating Board
  - Post-secondary education

APPENDIX C

OREGON BENCHMARKS: PRIORITIES FOR THE WORKFORCE QUALITY COUNCIL

BENCHMARK	1990	1992	1995	2000	2010
1. Percentage of high school students with significant involvement in professional technical programs	9%	9%	18%	35%	55%
2. Percentage of employer payroll dedicated to training and employment	1.5%		2%	2.5%	3%
3. Percentage of workforce that has received at least 20 hours of education related to work skills and knowledge within the past 12 months					
4. Percentage of displaced workers reemployed within 24 months and earning at least 90% of previous income		36%	60%	70%	75%
5. Percentage of adults who have completed at least one year of postsecondary education or training	56%	63%	65%	70%	75%
6. Percentage of 25-year-olds with a certificate granted in a non-baccalaureate education and training program		10%	15%	25%	40%
7. Prose Literacy (understands text information)					
a. Basic 21-25 years	77.7%				99%
All adults	78.0%				90%
b. Intermediate 21-25 years	38.0%				90%
All adults	41.1%				55%
c. Advanced 21-25 years	5.5%				50%
All adults	8.7%				25%
8. Document Literacy (can understand and use graphs, text, maps, etc.)					
a. Basic 21-25 years	78.2%				99%
All adults	76.1%				90%
b. Intermediate 21-25 years	37.1%				90%
All adults	35.5%				55%
c. Advanced 21-25 years	6.2%				50%
All adults	6.3%				25%

9. Quantitative Literacy (can understand math and apply it)						
a. Basic	**21-25 years**	76.9%				99%
All adults		80.0%				90%
b. Intermediate	**21-25 years**	27.5%				90%
All adults		39.0%				55%
c. Advanced	**21-25 years**	3.6%				50%
All adults		7.6%				25%
10. Percentage of adults who have completed a certified apprentice program			2.6%	3%	4%	6%

As can be seen from the table, baseline statistics were available for 1990 and/or 1992 for nine of the ten benchmarks. The desired level of achievement in 2010 for each of these benchmarks is presented as well. While the seven non-literacy benchmarks will be measured through annual or biennial surveys conducted by the state, data on the three literacy benchmarks will be gathered less often—perhaps every four to six years. This is because a statewide literacy survey is judged to be too expensive to be conducted more frequently. Consequently, no benchmarks for literacy were set for the years 1992, 1995, and 2000. Oregon will either participate in future national/state literacy surveys that will be conducted approximately every four years, or it will replicate the 1990 survey periodically.

Whether the targets will be met and the vision achieved does not hinge solely on how well public workforce agencies do their jobs. It depends also on the extent to which the private sector successfully promotes continued education and training for its workers.

APPENDIX D

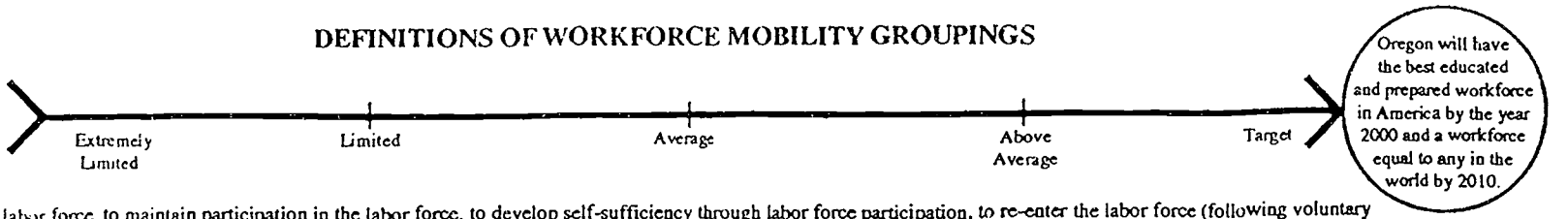
OREGON'S MOBILITY CONTINUUM

The interagency team that is working on performance measures is also working on a model for classifying people in regard to their abilities to enter, remain in, and progress through the labor force. Dubbed the "mobility continuum model," the goal is to eventually have a set of measures that will accurately reflect people's positions along a continuum from "extremely limited mobility" in the labor market (meaning that they have almost no skills or opportunities that would enable them to participate in the workforce) to "target mobility" (meaning that they possess all the characteristics needed to be part of a world-class workforce, including the ability to adapt to dramatic changes in labor force conditions.)

The idea is to use mobility-related measures in a pre- and post-test manner to document the main outcomes of workforce programs. This would allow fairer comparisons among programs serving populations that are at the same point on the continuum, and avoid unfair comparisons between programs serving fundamentally different populations.

Another implication of the mobility continuum is that it prompts the need to develop literacy measures to gauge progress (increased workforce mobility) at the higher end of the continuum, as well as at the lower end. Currently, using just the CASAS as the measure of performance, a participant's progress could not be tracked beyond a level of "average mobility."

DEFINITIONS OF WORKFORCE MOBILITY GROUPINGS



Mobility:
The ability to enter labor force, to maintain participation in the labor force, to develop self-sufficiency through labor force participation, to re-enter the labor force (following voluntary or involuntary unemployment), and to change occupations to reflect shifting employer requirements or personal needs. The purpose of employment and training services is to increase the ability of participants in order to ensure on-going, productive labor force participation leading to self-sufficiency. Mobility within the labor force is affected by both personal characteristics and community factors.

Extremely Limited Mobility:
Includes those who have never participated in the labor force and have few marketable skills. This group also will include individuals with many years of labor force participation in an occupation which is disappearing and/or which has provided few transferable skills. Includes those with limited basic skills (reading, writing, math, communication), those in isolated or depressed geographic areas, those with significant barriers to employer acceptance (i.e., conviction record, repeated job terminations, unacceptable attitude or appearance), and those with substantial physical, mental or emotional restrictions.

Limited Mobility:
Includes many of the "working poor"--individuals who hold jobs with low skill requirements and little opportunity to gain new skills or advancement. Neither the employer or the worker may be concerned about the limited mobility as long as employment conditions are stable. However, an employer's need to retrain "low tech" workers to operate advanced equipment in order to remain competitive will often highligh the hidden cost of having marginally trained workers. Conversely, dislocated timber workers are discovering the price of having basic skills and experience which are suited to a declining industry and do not readily transfer to other occupations. Mobility is also limited by family responsibilities, geographic location, educational background and available social services.

Average Mobility:
These individuals generally have good basic skills, an acceptable work history, and the resources and experience to maintain a structure in their lives which allows steady labor force participation. They most often have no dramatic barriers to employer acceptance and have experienced some expansion or upgrading in their skills while in the labor force. They primarily live in areas of occupational diversity. If they become unemployed, they generally re-enter the labor force quickly. Most often, their jobs enable them to be self-sufficient.

Above Average Mobility:
These individuals have education beyond high school, and will often have a two- or four-year degree or journeyman status. They will generally show a pattern of upward mobility in their jobs, and their work history will demonstrate significant marketable skills in their field. They usually participate in some form of on-going training. Most of this group will be computer literate and will have successfully dealt with major changes in job requirements or organization structure at least once during their careers.

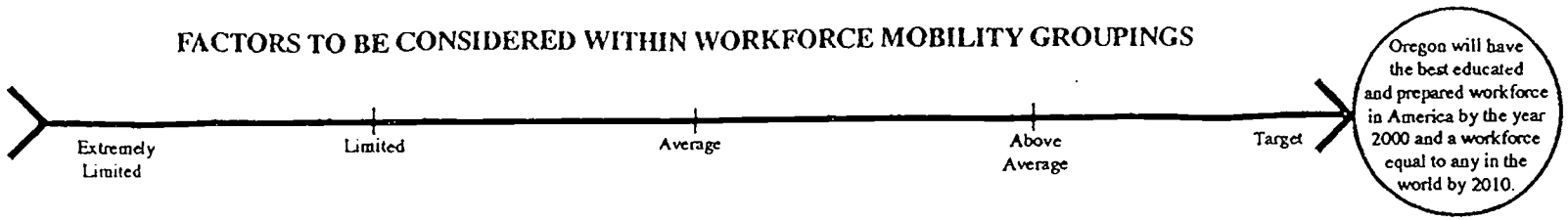
Target Mobility: (The best educated and prepared workforce in America by the year 2000 and equal to any in the world by 2010.)
Overwhelmingly, this group will have substantial education beyond high school. They will often have not only an undergraduate, professional or technical degree, but will also display a consistent pattern on continuing education and skill development (through personal initative and/or company assistance). These individuals will be proficient in utilizing computers and other technological aids; will demonstrate superior skill in dealing with significant personal and professional change; will be highly skilled at career development, including identifying and responding to workforce trends; and will be adept at working in a variety of organizational "cultures."

Please note: Descriptions of the mobility groupings do not include information on community mobility factors. These factors are just as critical as the personal mobility factors in reaching the target of the best educated and trained workforce in the world by the year 2010. However, there is currently much less information available to define precisely how communities and organizations can structure themselves to ensure the most productive workforce possible.

FACTORS ACROSS WORKFORCE MOBILITY CONTINUUM

Extremely Limited	Limited	Average	Above Average	Target
Personal Mobility Factors				
Very low basic skills	Low to marginal basic skills	Good basic skills	Solid basic skills	Excellent basic skills & advanced skills
Limited or no work history	Limited work history or extensive, narrow work history	Diverse, transferable work skills	Diverse skills with high transferability	Diverse skills, ongoing development, increasing marketability
Numerous physical, mental, emotional or social limitations	Some physical, mental, emotional or social limitations	Few physical, mental, emotional or social limitations	Have balanced packaged of physical, mental, emotional or social limitations	Uses resources to lessen impact of physical, mental, emotional or social limitations
Few physical, mental, emotional or social strengths	Some physical, mental, emotional or social strengths	A variety of physical, mental, emotional or social strengths	A variety of physical, mental, emotional or social strengths	Sense of interrelation of physical, mental, emotional or social strengths to job mobility
Significant family needs/extremely limited family resources (social & economic)	Relatively stable family needs/fairly limited family resources	Stable family/support system	Stable family/support system and family resources to support development	Stable family and support system; knowledge of resources and networks
No training opportunities	Extremely limited training opportunities	Ongoing, limited participation in education and training	Ongoing, full participation in education and training	Ongoing, extensive participation in education and training
Extensive dependence on government services	Sporadic dependence on government services	Little dependence on government services	High degree of self-sufficiency	Self-sufficient
Community Mobility Factors				
Isolated or depressed geographic area	Some occupational diversity in geographic area	Significant occupational diversity in geographic area	Support for a public policy of occupational diversity in community	Support for significant occupational diversity through strategic planning in geographic area
No dependent care available	Limited options for dependent care	Variety of dependent care options	Some dependent care available with some degree of accessibility	Dependent care available and accessible
No employer ability to accommodate or correct/accept-deficiencies	Limited employer involvement in addressing training & other needs	Recognition of employer involvement in addressing training and other needs	Significant employer involvement in addressing training and other needs	Significant & high quality employer involvement in addressing training and other needs
Pronounced employment discrimination and limiting employer bias	Some employment discrimination and limiting employer bias.	Limited employment discrimination and generally adaptive employer policies	Positive affirmative action program and adaptive employer policies	Aggressive affirmative action program and highly adaptive employer practices
No access to education/training	Limited access to education and training	Variety of education and training available	Enhanced selection and flexible delivery of education and training, and some incentives for participation	Extremely diverse selection & flexible delivery of education and training and incentives for participation

FACTORS TO BE CONSIDERED WITHIN WORKFORCE MOBILITY GROUPINGS



PERSONAL MOBILITY FACTORS:

• Diversity of Skills	The range of skills an individual has and the variety of occupational options created by the diversity of their skills. Increased diversity = Increased options
• Basic Skill Level	Reading, writing, math, interpersonal communication, problem solving, and "life" skills.
• Educational Background	Years of education, type of education, impact of education, appropriateness of education to labor market needs/personal circumstances, transferability of education.
• Physical/Mental/Emotional/ Social Limitations	Traits and characteristics which restrict various activities and which require accommodation on the part of the employer beyond that generally provided. These include physical or mental disabilities, inability to change, drug and alcohol involvement, low self-esteem, mental illness, immaturity, etc.).
• Physical/Mental/Emotional/ Social Strengths	Traits and characteristics which are viewed favorably by employers, may vary greatly by occupation/industry. Includes excellent health; physical strength; ability to learn quickly; good memory; cheerful and willing attitude; persistence; ability to work as a team member.
• Work History	Number and types of jobs, past performance and work behaviors, type of industry/occupation, salary history, etc.
• Family Characteristics	Number and ages of family members, needs of dependent family members, economic resources, support/stability available from family, family history of labor force participation, value placed on work and education, etc.

COMMUNITY MOBILITY FACTORS:

• Occupational Diversity	The range of occupations available in a geographic area.
• Dependent Care Options	Availability of & access to dependent care for children and other dependent family members.
• Housing	Housing options which support labor force participation.
• Transportation	The range of transportation options which allow access to the workplace.
• Services to Address Physical/Mental/Emotional/ Social Limitations	General availability of social services, mental health & drug & alcohol treatment, accessible buildings & streets for individuals with disabilities, groups & organizations (AA, religious & civic groups, Parents Without Partners, etc.) meeting emotional and social needs.
• Education/Training	Pertains to the range of education & training to which individual has access; relationship of training to current labor market needs & long-term trends, & the quality of education.
• General Economic Health	Overall economic health as measured by traditional standards (unemployment rate, interest rate, businesses start-ups & failures, profitability) as well as emerging indicators (ratio of part-time to full-time employment, increase in temporary and leased employees).
• Employer Bias	Attitudes and practices related to sex, race, age, religion, ethnic background, etc. May also include bias for traditional employer/employee relationships & management structures which may not be adaptive to current conditions.
• Employer Services Related to Mobility	Services which support increased mobility include: cafeteria style benefit plans, flexible scheduling, access for those with disabilities; and significant training for all workers.
• Health Care	Availability of & access to basic health care services.

APPENDIX E

EXAMPLES OF MEASURES DEVELOPED BY THE TEXAS WORKGROUP

Measures for External Reporting	Measures for Management Purposes
Adult Secondary Education	**Adult Secondary Education**
1. No. of JOBS clients enrolled in ASE during the fiscal year.	1. Percent of JOBS clients enrolled in ASE completing an adult secondary education level.
2. No. of JOBS clients in ASE who obtain a high school equivalency certificate during the fiscal year.	2. Percent of JOBS clients enrolled in ASE making progress in ASE, as measured by demonstration or progress, who are enrolled in ASE at the end of the fiscal year.
3. Percent of JOBS clients in ASE successfully achieving a high school equivalency certificate. (Measure 2/ Measure 1)	3. Percent of JOBS clients enrolled in ASE separated before the completion of a level who made progress within the level, as measured by demonstration and progress.
	4. Percent of JOBS clients enrolled in ASE separated before progress is made.
	5. Percent of scheduled hours of ASE attended for each JOBS client in ASE during the fiscal year.
	6. Average hours of attendance per ASE JOBS client during the fiscal year.
	7. Percent of JOBS clients in ASE referred to take the equivalency examination who achieve a high school equivalency certificate.

Case Management 1. Percent of IOBS clients enrolled in a component of the JOBS program who successfully complete the component. 2. No. of JOBS clients enrolled in a component who successfully complete the component. 3. No. of JOBS clients who move from one component to another. 4. Number of cases that are effectively directed (meet Quality Assurance standards).	**Case Management** 1. Percent of JOBS clients who successfully participate in JOBS components as measured by completion of 75% of their scheduled hours.
Individual Job Search 1. No. of JOBS clients who participate in Job Search during the fiscal year. 2. No. of JOBS clients in the fiscal year who become employed during (or within 30 days of completion of) the Job Search component. 3. Percent of JOBS clients who become employed during or just after completion of the Job Search component. (Measure 2/Measure 1) 4. Percent of JOBS clients enrolled in Job Search who were placed in employment by during the fiscal year. 5. Average wage at placement of Job Search participants during the fiscal year.	**Individual Job Search** 1. Percent of JOBS clients enrolled in Job Search who make progress in the activity as measured by meeting standards of conduct for job search.

APPENDIX F

DEVELOPING PERFORMANCE MEASURES

The Experience in Oregon

In order to establish accountability for state workforce programs, the Workforce Quality Council created a performance measurement committee composed of representatives from the ten agencies or divisions involved in workforce development. Programs operated by these agencies serve welfare recipients, displaced workers, potential labor force entrants, prisoners, underskilled workers, and others. The mission of this group is to recommend and/or develop a set of standard performance measures to be used by all workforce agencies. By applying the same measures to all public programs, their effectiveness can be more easily monitored and compared. Such a system would also provide the Council with information needed to guide policy and budgeting decisions.

Oregon has a well-defined vision of what its workforce should look like in 2010 and has developed statewide benchmarks for measuring progress toward this end. In addition, the Council has a specific mission to make the state's workforce development agencies accountable for results. The vision statement and benchmarks and, most particularly, the Council's stated mission defined the parameters within which the team has developed its initial set of program performance measures.

Their selected measures focus foremost on client outcomes — such as job placement and retention rates, wages, and skill gains — but also include information on program costs and customer satisfaction. The team notes that these are interim measures to be used until more refined measures can be developed. The hope is to produce a set of measures that can track an individual's progress along the workforce mobility continuum (A description of this continuum is provided in Appendix D). This means being able to measure improvements in literacy and job skills not only for people at the lower end of the continuum (those with few skills and little or no work experience), but also for people functioning at higher levels (such as workers whose current skills have become technologically obsolete).

In the long run, the plan in Oregon is to develop, test, and implement a set of fairly sophisticated outcome-based measures. In the short-term, the group was asked to propose a set of interim measures that could be used while the others were in development. Ideally, the proposed set of measures would meet the following criteria:

- The measures are sensible and fair.
- They are consistent with the vision and goals of the Workforce Quality Council.
- The measures are useful to the Council for evaluative, planning, and policy decisions.
- They evaluate return on investment.

The interagency group proposed and the Council adopted a set of five interim measures. The measures selected were: wages at job placement and follow-up, job retention, cost per placement, increased skill level, and customer satisfaction.

Placement wage information would come from program records. Information on subsequent wages and job retention for up to eight quarters after placement would be obtained by matching agency placement records with records in the state unemployment insurance tax files. The reasons for using the UI files for follow-up were that the information was readily available, it was relatively inexpensive to gather, and it was in a standard format. The major drawback to this method is that not all workers are covered in the UI file records, since certain classes of workers are exempt.

Cost per placement was included in the set of measures because it would help the Council evaluate return on investment. The group noted, however, that there are several potentially serious problems in measuring and reporting program costs. One is the lack of consistent definitions across agencies on what is a cost. The group found that agencies varied widely in what they currently included in their cost calculations. Considerable work would need to be done to standardize cost estimates across agencies. Another problem is that it may cost more to serve certain types of clients (e.g., those with fewer skills, greater needs, or more difficult personal circumstances). This makes it difficult to make fair cost-per-placement comparisons when programs are serving essentially different clientele. Finally, the group noted that including cost measures may serve to drive the program in unintended directions. If cost per placement becomes the paramount concern of decisionmakers, then "penny wise, but pound foolish" decisions may be made.

In terms of increased skill levels, the group recommended for the moment only measuring gains in literacy skills. The recommended method for doing this was by administering the Comprehensive Assessment System of Adult Skills (CASAS) to all clients in programs such as JTPA, adult basic education, or corrections, who don't have evidence of successful completion of post-secondary education or training. The CASAS would be administered at intake and following literacy training for those who were referred for such training. Gains in literacy skills as indicated by pre- and post-program test scores would be reported. If an individual completed multiple literacy training programs and was tested more than once, the difference between the initial and highest CASAS scores would be reported. The CASAS was selected as the measurement instrument of choice because of its ease of administration and relatively low cost, and because it has been used extensively by agencies in Oregon.

Customer satisfaction was the thorniest issue for the group. It was recommended that programs measure whether program participants and their employers were satisfied with the services and products they received. The rationale for including this measure was that programs should be customer driven, and it is always important to determine whether programs meet customer expectations. However, the group identified several potential problems in including satisfaction in the standard set of measures. First, they found it difficult to define satisfaction clearly. Second, information on satisfaction was not already being gathered; new scales would need to be developed, verified, and implemented for both participants and employers. Finally, the fiscal impact on the programs for this additional data collection was unknown.

Currently, the agencies are still in the beginning stages of implementing this set of standard measures. One early finding emerged, however, when cost-per-placement information was added in the last round of reports to the legislature. Committee members found that they must be as diligent in standardizing measures on the input side of the process (costs, client characteristics, program effort, etc.) as they are in standardizing outcome measures. When legislators were given cost-per-placement information on programs operated by different agencies — but without sufficient background information on who was served or how they were served — they sometimes reached the wrong conclusions about the efficacy of particular programs. For example, a job search program in one agency that serves welfare recipients who have no previous work experience should be expected to have a higher cost per placement than a program also called job search in another agency that serves more skilled and experienced clients.

As one committee member who was interviewed for this study commented, "There is a tendency for legislators and others to try to compare apples and oranges using the same cost-per-placement data. Sometimes they will even try to combine data from fundamentally different programs to determine the average cost per placement for programs that only sound like they are doing the same thing."

APPENDIX G

DEVELOPING PERFORMANCE MEASURES

The Experience in Texas

Although major client outcomes, such as job placement rates and wages, were already established for the JOBS program through the performance budgeting process, the legislature added a rider to the most recent appropriations bill directing the agencies to develop "intermediate outcome" measures that indicate whether JOBS clients are making adequate progress in the various components of the program.

Like Oregon, the Texas team has developed and is testing an initial set of measures that it intends to refine and improve over time. Unlike Oregon, the team is not attempting to standardize these measures across agencies — or even within agencies in some instances. What is standardized is the reporting of the measures: an agency simply reports "yes" or "no" to indicate whether each client is making adequate progress in its component of the JOBS program. The interagency team must agree in advance, however, that the measures of progress used by each agency are reasonable indicators of intermediate client outcomes.

Rider 33 was intended only to spur the development of *intermediate* outcome measures. The main goals, objectives, and outcomes for the various components of the JOBS program had already been defined through the performance budgeting process. Data collection and reporting systems were already in place, both for the performance budgeting process and for federal reporting on JOBS. In addition, the JOBS Interagency Planning Group, composed of senior officials from all of the state agencies involved in the JOBS program, had been in operation for several years. This group, as the state coordinating body for JOBS, had already developed a means for sharing information across agencies while ensuring the confidentiality of client records.

Against this backdrop, Rider 33 can be viewed as an attempt by the legislature to correct and improve the existing outcome-based data collection and reporting system associated with JOBS. It was an honest admission that: "Hey, we didn't get it all right the first time. Let's go back and add some things that are missing." The Rider 33 workgroup was able to accomplish what it did in the time allotted in part because this was essentially an extension of established systems; it did not require the creation of a whole new interagency data system nor the forging of new interagency partnerships. Without this solid base, the process would have taken much longer.

A review of the 78 measures developed through the Rider 33 workgroup efforts reveals that many are clearly defined measures of interim outcomes — e.g., "Percent of adult secondary education JOBS clients successfully achieving a high school equivalency certificate." (Measures for several of the JOBS components appear in Appendix E). In other instances, the intermediate outcome measures are stated almost tautologically: "The percent of JOBS clients enrolled in [name of activity] who make progress in the activity," where the term "progress" is later defined by a phrase such as "as measured by demonstration and performance." Further review of the list reveals that many of the measures are not really outcome measures, even of the intermediate kind. There are lots of measures on the list in the form: "Number of JOBS clients who participate in

[name of activity] during the fiscal year." According to definitions used by the Texas legislative budget office, these are really *output measures*, which are defined as "The quantity of agency workload and work product as it pursues its strategies." They are measures of the level of effort expended by the agency, not of outcomes (or even of progress toward outcomes) for the JOBS clients.

How did the workgroup end up with such a mixed list? The answer, interviews revealed, is that the list of measures reflects the results of reasoned decisions made by the group in response to (1) the particular constraints and opportunities encountered in each area for which they were developing measures, (2) the deadlines imposed by Rider 33, and (3) the expressed needs of legislative and state auditor's staff for additional kinds of information. A brief review of the process used by the workgroup to select the measures will show how decisions made at various points in the process culminated in the current list.

First, each agency was asked to identify or propose intermediate outcome measures for every JOBS component it operated. Ideal measures would meet the following criteria:

(1) validity — actually measures client progress in the component;

(2) meaningfulness — the measure makes sense to others when reported to them;

(3) relative ease of collection — gathering the data does not impose an excessive burden on service delivery staff; and

(4) availability — preferably, the measure is already being collected.

Workgroup members were able to find existing measures for many of the components. If none existed for a particular component, the agency had to create one. In cases where a training program was fairly standardized across the state, this was not a difficult task. Providers could be instructed, for example, to track and report as a measure of progress the number of training modules a client completed. Sometimes an agency decided to create a new assessment instrument that could be administered when clients enrolled in a program and then again several weeks later to gauge progress. In areas such as English as a second language and adult basic education, where service delivery is highly individualized, it was decided that progress would be measured by an individual's demonstrated performance judged against criteria determined by the local program.

The auditor's staff also recommended including a measure of the quality or "goodness" of case management in JOBS. After much deliberation, the workgroup decided to assess this crucial aspect of the program in two ways:

1. Tracking for all components combined the percentage of clients who make progress within a component, complete a component, and/or move from one component to another; and

2. Conducting annual quality assurance audits of a sample of representative cases.

The audits rate each case record on the appropriateness of the case plan for meeting the client's needs, whether barriers to client progress are removed, whether the plan is monitored for effectiveness, and whether any lack of progress is quickly addressed by the case manager.

Second, the workgroup examined the proposed measures and approved a minimum standard of progress in each. Usually the group approved the standard proposed by the agency responsible for the particular component. Sometimes, however, the workgroup as a whole debated and decided how much change constituted progress in a particular component. It decided, for example, that clients in job search who meet three of five criteria for success by the third week of the program should be counted as making progress.

The group tried to standardized the presentation of the measures to the extent possible. The state fiscal year was selected as the standard reporting period, and the percent of enrolled clients who were making progress in a component was the preferred statistic. Most measures could then be presented in the form: "Percent of JOBS clients enrolled in [a particular activity] who [completed beginning level; scored higher on a second assessment; etc.] during the fiscal year."

Third, the proposed measures were submitted for review and approval to the legislative budget office and the state auditor's office. At this point, most of the current input measures were added to the list because these entities indicated they also wanted the numbers on which the percentages were based to be reported.

Finally, the legislative and audit staff divided the list of approved measures into two categories. Measures for external reporting would be reported to the legislature and auditor, and measures for management purposes would be used solely by the agencies administering the components of the JOBS program.

Workgroup members stressed in the interviews that the current list of intermediate outcome measures should be viewed as a work-in-progress. The implementation phase of gathering and reporting these measures has just begun, and they anticipate that this experience will undoubtedly point out needed changes. The first reports on these measures for external reporting purposes will be prepared by December 1994. The workgroup will then get feedback on the utility of the measures both for program management purposes within agencies and for JOBS program monitoring by the legislative budget office and state auditor's office. The intermediate outcomes and their measures will be revised in response to this feedback.

APPENDIX H

SHARED INFORMATION SYSTEMS IN OREGON

One of the chief goals of the Workforce Quality Council is to develop a fair and reliable way to determine the relative return on investment for the various services provided by the education and human resource agencies that comprise the state's workforce development system. Having every program use a standard format to report information on its participants and their outcomes would help the Council improve performance measurement, evaluation, and planning across agencies. It would also give the Council the information it needs to recommend where to invest state resources most effectively. If this information could be shared electronically across agency computers, it would allow the Council to see how people actually use the workforce development system. For instance: Do they use services from different agencies? If so, how many and in what sequence? Do they ever enroll in essentially the same services offered by different agencies?

With these goals and questions in mind, the Council directed the team of representatives from the ten workforce development agencies to identify strategies for promoting the electronic sharing of information among agencies. Their charge was to propose, test, and implement a shared information system that would:

1. Improve the collection and reporting of client profile data;

2. Enhance the state's ability to make informed policy decisions and improve programs;

3. Enhance the Council's ability to evaluate decisions, programs, and resource allocations;

4. Enhance the ability of participating agencies to collect and report program data in support of uniformly developed performance measures (such as the five interim measures discussed above); and

5. Improve the efficiency and cost effectiveness of service delivery to Oregon's citizens.

The interagency team began its work in 1992 by conducting an "application transfer study" in which focus groups of potential users of information from the electronic system described their information needs and the barriers they saw to sharing information. The groups then recommended methods for measuring performance and ranked alternative solutions for overcoming barriers. In all, 200 people at all program levels from the ten agencies were interviewed. The process was facilitated *pro bono* by IBM.

The basic form for the shared information system (SIS) emerged from this study and other research conducted by the interagency team. The decision was made to use the workforce mobility continuum model as the basis of the SIS. A standard set of client measures would be used to indicate where each program participant fell on the continuum in terms of education, job skills, and other factors related to workforce mobility. Program outcome measures, such as skill gains and job retention, would also be standardized. Ideally, each of the ten agencies would maintain this

information in their computer data bases, and the agency computers would be linked so that client records could be matched and information on people served by more than one agency could be shared.

A list of tangible and intangible benefits to be derived from such a system was developed by the committee in order to convince the 1993 legislature to authorize the establishment of the SIS. These were:

Tangible benefits — true potential savings or cost avoidances.

- reduction of duplications,
- streamlined, uniform program evaluation,
- increased efficiency and effectiveness of resource allocation, both within and across agencies.

Intangible benefits — enhancements to program operations and client services that tare not directly measurable.

- increased staff synergy leading to program improvements,
- better focus by workforce agencies on council goals and Oregon benchmarks
- increased visibility and accountability for workforce development services
- development of a common language (definitions and performance measures) for conducting analyses.

The committee decided to begin the process with five units that already shared a local area computer network: the Employment Division, Vocational Rehabilitation Division, Worker's Compensation Division, Adult and Family Services Division, and the Job Training Partnership Act Administration. The other agencies will be added, once the SIS is tested and operational. When work began on the system, two issues immediately came to the fore: confidentiality of client records and the ability of the participating agencies to adopt common data standards.

Working throughout 1993, the Oregon committee representing the five units involved in developing the SIS generated a proposed set of standard data elements to be used by all agencies for recording and reporting client information and outcomes. In all, seventy-five data elements were deemed crucial for creating a SIS based on the mobility continuum. The major categories of elements were:

- Record identification — agency, date, social security number, etc.;
- Client intake information — birth date, sex, county of residence, annual income, marital status, occupation, last grade completed, etc.;
- Program/services — services received, training and programs participated in, completion codes, costs for direct and supportive services, etc.;
- Job placement information — type of employment, wages, hours, etc.;

- Skill gains — improvements in reading and computational skills as measured by the CASAS;
- Satisfaction — client and employer satisfaction scores;
- Labor market experience — post-program quarterly wages and job retention, employer code, etc.

Many of these data elements were already being recorded by the agencies. The group found, however, that on average each agency would need to add 15 to 20 new elements to its current client data collection system. In addition, it was found that agencies often did not record the information they were already collecting in the same way. For instance, one agency might use "M" or "F" to record a participant's gender, while another used "1" or "2." One agency might record the highest grade completed in school, while another might report the number of years of education. To remedy this situation, the committee developed a the 120-page data dictionary that contained standard definitions for all of the proposed data elements in the SIS.

Several problems with this standardized approach to data collection emerged when the five pilot agencies began to implement the SIS. First, the funds appropriated by the legislature for the SIS — approximately $800,000 — only covered a portion of the costs for converting existing data systems and collecting additional information on each client. The agencies had not budgeted for the additional costs in their biennial appropriations requests. Second, it was found that some agencies would have more difficulty gathering all of the desired information than others. In these agencies, new client forms, intake and tracking procedures, and computer systems would need to be developed. Third, some of the requested data elements did not fulfill the operational and management needs for some agency programs. The elements were seen as perhaps useful to the Council for evaluation and planning purposes, but of little use to the agencies collecting the information. Therefore, these agencies had little incentive to invest their own funds in collecting the additional data.

Finally, there was the issue of data quality. To get high-quality, accurate data from the agencies requires both up front training of line workers and others who will record the data, as well as constant vigilance in data collection to ensure the completeness and reliability of the data. One of the committee members noted that the issue of data quality is particularly troublesome for elements that are not seen as useful to the agency and its workers: "If staff can't see how collecting a certain bit of information will be beneficial to them in their work—or to their managers—they will not care much whether the data [they report] is accurate or complete. Unless, of course, there are sanctions for poor reporting such as loss of program funds or a poor performance review for the worker."

Faced with these potential problems, the interagency group at the close of 1993 considered three options that would allow the five agencies to begin standardized data collection in 1994 and issue a required report to the legislature in 1995: (1) require all workforce development agencies to collect the complete set of SIS data on all clients; (2) collect complete SIS information only on a sample of clients from each agency; (3) develop agency-specific models that encourage, but do not require, complete reporting of all elements by the agencies. The main advantages to the latter two

options were cost savings in data collection and the ability to get much, but not all, of the SIS data in a standard form in the time allowed. The sampling option would ensure high quality data for sampled clients, but the reported results would be subject to some sampling error and would not be accurate below the regional level.

In the end, the Council settled on a compromise among the presented options. Agencies were permitted some latitude in selecting the elements they would report, but they were also encouraged to report as many as possible given their budgetary and workload constraints. Agencies were required to use the standardized format if they chose to report some data elements — notably client characteristics and program outcome and follow-up information — but other elements such as program cost measures were not standardized. For the most part, data were collected and reported on all program clients; however, for some elements, such as client and employer satisfaction, sampling methodologies were used.

The first reports from the SIS on the five pilot agencies will be presented to the legislature in 1995. The plan is to review the utility of the system and its initial reports, to make modifications as needed, and to begin working with the other five workforce agencies so that they can be part of the SIS in the next biennium.

An analyst at the Workforce Quality Council, when asked to assess the state of performance measurement for the workforce agencies, replied: "We have been at this in earnest since 1992. Our first reports come out next year. I think it will be 1997 before we have something in place in Oregon that we can be proud of. Realistically, the system won't be operating at the high level of quality we expect until the year 2000. It really takes about ten years to fully develop and implement a system such as this.

APPENDIX I

SHARED INFORMATION SYSTEMS IN TEXAS

The Texas Rider 33 work effort was not evenly distributed among the agencies. Those that operated more JOBS components or had to create new data collection methodologies for their components naturally had more work to do than the others. The Texas Department of Human Services (TDHS), as the coordinating agency for JOBS, and as the repository and compiler of data from all the other agencies, shouldered the lion's share of the work.

In September, the participating agencies began implementing the system for collecting and reporting intermediate outcome data on the JOBS program. The first report to the legislative budget office and the state auditor is scheduled for December 1994. Just because the data collection system is up and running, however, does not mean that the work of the Rider 33 workgroup is done — far from it. Group members estimate that it will take at least two years of ("not quite so intense") work to "shake down" the system and bring all the parts up to speed.

Data on the intermediate outcomes for JOBS clients is generated by the case managers, trainers, employment specialists, instructors, teachers, and others who provide direct services to the clients. To minimize the data collection burden for these people, no new reporting instruments were added for any of the components. Instead, existing forms from several agencies (or clients' computer data screens if electronic forms were used) were modified slightly to allow the service providers to simply check whether a client is making adequate progress according to the standards set by the workgroup for that component. This information, along with other client-specific data normally reported by the agency, is submitted electronically to TDHS each month. Workgroup members noted that they were not able to reduce or eliminate the collection of any process-oriented data elements since most are required for federal reporting on JOBS.

Generally, the data submitted to TDHS will be compiled monthly and returned to the sending agency in the form of management reports. These reports will provide information on client progress and outcomes summarized to the local office or provider level. Workgroup members from other agencies who were interviewed for this study said the reports from TDHS will provide powerful management tools that were not available prior to the Rider 33 collaboration. Although this type of reporting has the potential to allow managers to set performance standards for regional and local offices or providers, the agencies intend to move slowly on this, preferring to first establish a solid baseline of performance measurement. Meanwhile, the workgroup will seek to determine the usefulness of the reports to managers and will modify them if necessary.

The monthly intermediate outcome data will be aggregated to provide the annual reports required by the legislative budget office. As with the other reports, the content (and possibly the format) of the annual legislative report will evolve over time.

Although it is now in the implementation phase, the group cannot dust its hands and declare the measurement development process over. Many of the measures initially developed need to be carefully validated and revised if necessary. The workgroup will have many

opportunities to improve the measures, reports, and procedures over the next couple of years as the implementation phase moves into high gear.

Texas is also developing a computer data base to combine client demographic and outcome information across agencies. The major difference between it and Oregon's shared information system is that the Texas system is not only used to provide annual reports to oversight entities, but it is also designed to provide program management information to the agencies and local program operators. This means producing detailed monthly reports for each agency that can be used to promote continuous improvement in program operations.

The final notable difference between the work undertaken in Oregon and that in Texas concerns their relative timeframes for developing measurement and reporting systems. In Oregon, it was assumed from the beginning that creating an adequate measurement system would require a considerable investment of resources over a period of years. Work has been ongoing for several years and is likely to continue for the foreseeable future. In Texas, where the scope of work is not as ambitious, the interagency team had less than a year to design its system before beginning data collection. Yet, here too they acknowledge that it will require at least two more years of work before the system is fully developed and operating adequately.

NATIONAL INSTITUE FOR LITERACY
800 Connecticut Avenue, N.W., Suite 200
Washington, DC 20006

202/632-1500 *202/632-1512 fax*